Stochastic Frontiers and Asymmetric Information Models

Philippe Gagnepain, Marc Ivaldi

Stochastic Frontiers and Asymmetric Information Models

Philippe Gagnepain
Departamento de Economía, Universidad Carlos III, Madrid, Spain

Marc Ivaldi[*]
GREMAQ - IDEI and Ecole des Hautes Etudes en Sciences Sociales, Toulouse, France

June 2001
This version: March 2002

Correspondence

Philippe Gagnepain
Department of Economics
Universidad Carlos III de Madrid
C/ Madrid, 126; 28903 Getafe, Spain
Tel: (+34) 91 624 5732
Fax: (+34) 91 624 9875
Email: philippe@eco.uc3m.es

[*] The authors are grateful to the Commissariat Général du Plan for its financial support. They benefited from comments of participants of various seminars or conferences. Under a slightly different title, this paper has been presented at The ASSET Conference, Marseille, 1997, Le Séminaire Jeune Gremaq, Toulouse, 1998, The 8th international Conference on Panel Data, Göteborg, 1998.

Abstract

This article is an attempt to shed light on the specification and identification of inefficiency in stochastic frontiers. We consider the case of a regulated firm or industry. Applying a simple principal-agent framework that accounts for informational asymmetries to this context, we derive the associated production and cost frontiers. Noticeably this approach yields a decomposition of inefficiency into two components: The first component is a pure random term while the second component depends on the unobservable actions taken by the agent (the firm). This result provides a theoretical foundation to the usual specification applied in the literature on stochastic frontiers. An application to a panel data set of French urban transport networks serves as an illustration.

Keywords: Costs and production stochastic frontiers; asymmetric information; technical inefficiency; effort; regulation; test of specification; urban transport.

Invoking the literature on asymmetric information models, this article reexamines the foundation of the econometrics of stochastic production or cost frontiers. Originally proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977), a frontier model consists of a function of the usual regression type with an error term comprising two parts. The first part corresponds to the usual white noise process. The second part represents global inefficiency. As it is well established in the literature, global inefficiency of individual sample firms can be predicted on the basis of cross-sections or panel data sets on these firms. The method based on cross-sectional data suffers from one serious difficulty: The estimation procedure must assume that inefficiency is independent of regressors. This might be incorrect since input and output quantities are together determined at the equilibrium and since firms may know something about their level of inefficiency when they choose inputs quantities. This assumption is potentially avoidable if one has panel data sets. (See Schmidt and Sickles, 1984. Cornwell, Schmidt and Sickles, 1990, and Ivaldi, Perrigne and Simioni, 1994, propose models where inefficiency varies over time.)

Now consider the case of a regulated monopoly. Through the window of a simple principal-agent model, informational asymmetries between the regulator and the monopoly affect the production process. This article provides an example to show that the production frontier of this regulated monopoly involves a global inefficiency term that comprises two terms. The first component is a purely exogenous random term; the second one is endogenous in the sense that it depends on the producer’s actions and hence on observable variables in an indirect and imbricated way. This decomposition of global inefficiency resembles the specification commonly used in the literature on stochastic frontiers. However, it is here endogenously derived while, in the tradition of stochastic frontiers, it is imposed in ad hoc way.

This resemblance finds its source in the economic literature. On one side, a tradition initiated by Leibenstein (1966) motivates the specification of stochastic frontiers. Without referring explicitly to the notion of frontier, Leibenstein clearly mentions the existence of a global inefficiency that depends on the will of managers and workers in a production process. A low powered incentive environment due to a lack of productivity of working agents induces the inefficiency, while appropriate incentives can lead to significant operating cost reductions.[1]

On the other side, since the emergence of the new theory of regulation, economists admit that, in industries where a producer is regulated by an authority, the principal-agent relationship is at the core of the question of assessing the performance of a firm. (See Loeb and Magat, 1979, Baron and Myerson 1982 and Laffont and Tirole 1986.) Hence, technical inefficiency and effort are two unobservable parameters, which characterize the informational asymmetries and define the source of global inefficiency.

In this perspective, asymmetric information models provide a relevant framework for the analysis of production and cost frontiers.[2] In addition, because such models are able to elicit the structural relationship between the observable variables and the inefficiency term, they directly provide a way to deal with the endogeneity of the inefficiency term, i.e., with the stumbling block of the econometrics of stochastic frontiers. However there is a cost to pay. Our example of principal-agent model shows that data on costs are now required to fully identify the endogeneity of the inefficiency term in a production frontier.

The analysis presented below is based on an example of regulatory structure drawn from the French urban transport industry. Here a regulated producer provides a single good or service. The regulator sets the level of supply to be offered by the operator. The former ignores the technical skills of the latter and its investment in cost reducing activities. Hence, the regulator does not know in advance how far the actual level of production deviates from the theoretical frontier. Section 1 presents the structural cost and production frontiers to be estimated using a Cobb-Douglas technology. Although this specification may not be the best choice for achieving a good approximation of a real world cost function, we choose it for its tractability and for ease of exposition and interpretation. Section 2 is devoted to an application of the derived cost model to the case of the French urban transport industry. Section 3 concludes.

1. The structural approach

The aim of this section is to construct the structural cost and production frontiers of a regulated monopoly. First, preliminary frontiers, that are conditional on effort, are obtained. Second, using the regulatory constraints, we construct structural frontiers to be estimated.

1.1 ***A three-components structure***

A standard specification of the stochastic production function is

(1) $$\ln Y = \ln f(X,\alpha) - \theta + \varepsilon_y \,.$$

Here, the levels of output and inputs are given by Y and X respectively. The parameter θ is a one sided, non-negative term, expressing global inefficiency, i.e., unobservable effects such as productive ability and managerial effort that have bounded effects on outputs. The second one, ε, is a two sided statistical noise which accounts for measurement errors. In this expression, the unobservable part is then

(2) $$u = \varepsilon - \theta \,.$$

We decompose now the global inefficiency term θ into an exogenous and an endogenous component. Hence, the unobservable part is

(3) $$u = g(e - \theta) + \varepsilon \,,$$

where $\theta \geq 0$ is regarded now as the exogenous technical inefficiency, $e \geq 0$ is the effort of productivity exerted by the firm and ε is the error term presented above. Technical inefficiency represents the amount of knowledge, the experience of the monopolist and its ability to associate inputs with the technology available. The cost reducing effort is exerted in order to reduce the technical inefficiency θ. The function $g(e-\theta)$ provides then a measure of the global inefficiency. The nature of this function depends on the source of asymmetric information in the production process. We assume that $g(.)$ is strictly increasing with θ and strictly decreasing with *e*. A measure of the total distortion under the production frontier is given by the ratio

(4) $$Y/[f(X,\alpha)\exp(\varepsilon)] = \exp[g(e-\theta)].$$

1.2. ***Preliminary frontier***

Production frontier

The regulator does not know neither θ nor *e*. We assume that the inefficiency, as well as the asymmetric information faced by the regulator, interferes with the utilization of some inputs only. In order to provide the level of output *Y*, a regulated monopolist requires quantities of input X_j, $j = 1,..n$, from a set *X* of inputs. The utilization and the management of a part of the set of inputs are affected by technical inefficiency θ. This leads to an over-consumption of one or several inputs. By exerting a significant level of effort, the monopoly can reduce this excess of factors demand. Let *i* be the factor affected by technical inefficiency. We assume that $i = 1,...,l$ and $l < n$. Thus, only a maximum of $n-1$ inputs can be affected.

Then we distinguish X_i from X_i^*. On one hand, X_i is the physical amount of input *i* used by the producer in the process. This amount is observable by the regulator. On the other hand, X_i^* is the actual level of input *i* which is not observable by the regulator. Hence, operating costs depend on the quantity X_i whereas the actual level of production depends on X_i^*. The relationship between observed and actual quantities is given by

(5) $$X_i^* = \frac{X_i}{\exp(\theta - e)}.$$

Thus, if $\theta - e$ equals 0, all units of input *i* are efficient and fully contribute to the process. Note that θ is expected to be greater than *e*. However, we do not impose this constraint when we estimate the model.

Consider now a Cobb-Douglas technology. In the literature on frontiers, the usual production function is

(6) $$Y = A \prod_{j=1}^{n} X_j^{*\alpha_j} k^{\alpha_k} \exp[-\theta + \varepsilon_Y],$$

where *k* stands for capital and A, α_k and the α_j's are parameters describing the technology. Capital is considered as a fixed input. Here, given that only *l* inputs are affected by inefficiencies, we rather consider a stochastic production frontier specified as follows:[3]

(7) $$Y = A \prod_{j=1}^{n} X_j^{\alpha_j} k^{\alpha_k} \exp\left[(e - \theta) \sum_{i=1}^{l} \alpha_i + \varepsilon_Y \right].$$

The global inefficiency $e - \theta$ is then weighted by the sum $\sum_i^l \alpha_i$, which increases with the number of inputs affected by productive inefficiencies.

Cost frontier

We turn to the construction of the dual cost frontier. Assume that the producer seeks to minimize the cost C of producing its desired rate of output Y under technical inefficiency. The associated allocation of inputs sets the factor demand $X_j,\ j=1,...n$. The program of the producer is then

(8) $$\min_{X_j} \sum_{j=l}^{n} w_j X_j \ ,$$

under the technological constraint

(9) $$Y = A\prod_{j=1}^{n} X_j^{\alpha_j} k^{\alpha_k} \exp\left[\left(e-\theta\right)\sum_{i=1}^{l}\alpha_i\right],$$

where w_j is the price of input *j*. The demand for factor *j* associated with this program is

(10) $$\ln X_j = \ln\kappa_j + \ln\left\{\prod_{m=1}^{n}\frac{w_m^{\alpha_m/r}}{w_j}\right\} + \frac{1}{r}\ln Y - \frac{\alpha_k}{r}\ln k + \frac{\sum_{i=1}^{l}\alpha_i}{r}\left(\theta - e\right),$$

where

(11) $$r = \sum_{j=1}^{n}\alpha_j \ ,$$

and

(12) $$\kappa_j = \alpha_j\left(A\prod_{j=1}^{n}\alpha_j^{\alpha_j}\right)^{-1/r} .$$

The excess of factor demand above its frontier prevents the producer from reaching the theoretical level of production. Such a distortion of factor demands leads to a rise of operating costs. Note that *r* provides a measure of the returns to scale in the production process. As the returns to scale increase, the factor demand distortion above the frontier reduces.

From (10), the stochastic cost frontier is

$$\ln C = \mathrm{K} + \sum_{j=1}^{n} \frac{\alpha_j}{r} \ln w_j + \frac{1}{r} \ln Y - \frac{\alpha_k}{r} \ln k + \sum_{i=1}^{l} \frac{\alpha_i}{r} (\theta - e) + \varepsilon_C \,, \tag{13}$$

where

$$\mathrm{K} = \ln\left(\sum_{j=1}^{n} \kappa_j \right). \tag{14}$$

Note that the error term ε_C in Equation (13) accounts for measurement errors. The term $(\theta - e)\sum_{i=1}^{l} \alpha_i / r$ in Equation (13) represents global inefficiency and $\exp\left((\theta - e)\sum_{i=1}^{l} \alpha_i / r\right)$ is the total cost distortion above the frontier. Global inefficiency is less significant when the industry enjoys large returns to scale.

Finally, note that Equation (13) can be rewritten as

$$\ln C = \mathrm{K} + \sum_{i=1}^{l} \frac{\alpha_i}{r} \left[\ln w_i + (\theta - e)\right] + \sum_{h=n-l}^{n} \frac{\alpha_j}{r} \ln w_h + \frac{1}{r} \ln Y - \frac{\alpha_k}{r} \ln k + \varepsilon_C \,. \tag{15}$$

Recall at this stage that there are *n* inputs in the process. Among them, *l* are affected by inefficiencies; hence *n-l* are not affected and fully participate to the production process. Thus, index *j* refers to any input involved in the process. Index *i* refers to any inefficient input; index *h* refers to any efficient input. From Equation (15), the dual cost function is of the form

(16) $$\ln C = C\left(\mathrm{K}, w_i^*, w_h, Y, k, \alpha, \varepsilon\right),$$

where

(17) $$w_i^* = w_i \exp\left(\theta - e\right).$$

We know that, to be allocatively efficient, the production organization must choose a combination of inputs that recognizes relative input prices and marginal products. Thus, the production process must follow the rule

(18) $$f_j\left(X_j\right) / f_{j'}\left(X_{j'}\right) = w_j / w_{j'}, \quad \forall j \neq j'.$$

Equations (16) and (17) show that the technical inefficiency θ affects the equality (18) since the input price w_i is distorted upward. The equality (18) is then no longer valid. Indeed,

(19) $$w_h / w_i > w_h / w_i^*, \quad \forall\, i = 1,\ldots,l, \quad \forall\, h = 1,\ldots,n-l.$$

This suggests that the technical inefficiency that affects l inputs leads to a general allocative inefficiency. There is a relative price distortion for inputs i. The producer minimizes costs by taking into account higher prices for these inputs. Such a general allocative inefficiency disappears when all the inputs are affected by the same amount of primal technical inefficiency, i.e., when $l = n$.

Both production and cost frontiers in Equations (7) and (13) respectively allow the econometrician to estimate a technology in a similar way. The choice between estimating one functional structure or the other depends upon exogeneity assumptions. A cost function should be rather considered if output quantities are exogenous. Nevertheless, in any case, both frontiers in

Equations (7) and (13) are preliminary since the unobservable structure is partially endogenous. The cost reducing effort *e* is endogenous and depends on regulatory schemes set by the regulator.

It is considered that, in a regulated environment affected by informational asymmetries, the parameters θ and *e* are, at the same time, components of global inefficiency in the production process and source of asymmetric information in the contractual game between the producer (the agent) and the regulator. We turn now to the regulatory aspect of the problem.

1.3. ***Incentives***

Assume that the producer is a monopoly. This producer is regulated by an authority, which may provide transfers to the operator and raise taxes. Assume in addition that two types of regulatory contracts are used: cost-plus or fixed-price contracts. The regulator sets the levels of prices, quality, supply and capital. Under a cost-plus contract, the regulator receives the commercial revenue *R* and pays the operator the operating costs *C* and a net transfer t_0. In a fixed-price regime, the operator bears all the risk on cost and revenue. It just receives a net transfer $t_0 = C_0 - R_0$ that ensures expected budget balance. Expected costs and revenue are C_0 and R_0 respectively.

Given the regulatory mechanisms, firms' pay-offs *U* are given by

(20) $$U = t_0 + \rho\left[R(Y) - C(Y, e, \theta)\right] - \psi(\mathrm{e}),$$

where $\rho = \{0,1\}$ when the operator faces a {*cost plus, fixed price*} type regime. Exerting effort is costly and leads to internal cost $\psi(e)$. A profit-maximizing operator determines the optimal effort level in Equation (20). The first order condition is

(21) $$\psi'(e) = -\rho C_e.$$

It states that the optimal effort level equalizes the marginal disutility of effort and the marginal costs savings under fixed-price regimes while it is supposed to be nil under cost-plus regimes. Let us define a specific functional form for the internal cost of effort:

(22)
$$\psi(e) = \exp(\tau e) - 1,$$

where $\tau > 0$. We assume $\psi'(0) = 0$. The first order condition (Equation 21) can then be expressed using the cost frontier (Equation 13) and the internal cost of effort (Equation 22). The level of endogenous effort exerted by the operator is

(23)
$$e = \rho \left[\frac{\mathrm{K} + \ln \sum_{i=1}^{l} \frac{\alpha_i}{r} + \sum_{j=1}^{n} \frac{\alpha_j}{r} \ln w_j + \frac{1}{r} \ln Y - \frac{\alpha_k}{r} \ln k + \sum_{i=1}^{l} \frac{\alpha_i}{r} \theta - \ln \tau}{\tau + \sum_{i=1}^{l} \alpha_i / r} \right].$$

The optimal effort level depends on inputs prices w_j, production level *Y*, capital stock *k*, the inefficiency θ and the technology available α. This shows why the term denoting global inefficiency in the preliminary frontiers is correlated with the regressors. The variables of the production frontier (Equation 7) and the cost frontier (Equation 13) appear in the expression of the effort at the optimum. Note that, in our example, this is true when the regulatory scheme is a fixed-price scheme only. In this case, the econometrician may have to take care of a serious problem of endogeneity when estimating a production or a cost frontier.

1.4. ***Identification***

The optimal effort level (Equation 23) is now reintroduced in the preliminary frontiers (Equations 7 and 13) in order to derive the final structural frontiers to be estimated. We make the following propositions:

PROPOSITION 1: *The structural production frontier is*

(24)
$$\ln Y = \rho\left\{ \mathrm{H}_Y + \sum_{j=1}^{n}\alpha_j \ln X_j + \alpha_k \ln k + \sum_{i=1}^{l}\frac{\alpha_i}{r\tau}\left[\sum_{j=1}^{n}\alpha_j \ln C_j\right] - \sum_{i=1}^{l}\alpha_i\theta \right\} +$$
$$(1-\rho)\left\{ \ln A + \sum_{j=1}^{n}\alpha_j \ln X_j + \alpha_k \ln k - \sum_{i=1}^{l}\alpha_i\theta \right\} + \varepsilon_Y,$$

with

(25)
$$\mathrm{H}_Y = \left[1 + \sum_{i=1}^{l}\frac{\alpha_i}{r\tau}\right]\ln A + \sum_{i=1}^{l}\frac{\alpha_i}{\tau}\left[\mathrm{K} + \ln\sum_{i=1}^{l}\frac{\alpha_i}{r} - \ln\tau\right],$$

(26)
$$C_j = L_j w_j .$$

This proposition motivates the following comments: When the effort of the producer is nil (under a cost-plus regime), no endogenous effect needs to be identified in the global inefficiency term. The structure of the frontier is similar to the one that can be found in the usual literature. Consider now the case of a producer regulated under a fixed-price regime. Since the producer is in this case residual claimant for cost overruns, he provides a strictly positive effort level. Hence, the global inefficiency is partly endogenous and needs to be explicitly expressed using the available ingredients of the model. This approach is however costly since data on input prices are necessary to fully disentangle the effects of moral hazard (the endogenous part of the global inefficiency term) from the effects of adverse selection (the exogenous part of the global inefficiency term). Indeed, from Equation (23), it can be seen that the optimal effort provided by the producer depends on the input prices w_j. Reintroducing this effort expression in the preliminary production frontier (Equation 7) leads to a functional form (Equation 24) that depends on both input quantities L_j and prices w_j. We turn now to Proposition 2:

PROPOSITION 2: *The structural cost frontier is*

(27)
$$\ln C = \rho\left\{\mathrm{H}_C + \xi\left[\sum_{j=1}^{n}\frac{\alpha_j}{r}\ln w_j + \frac{1}{r}\ln Y - \frac{\alpha_k}{r}\ln k + \sum_{i=1}^{l}\frac{\alpha_i}{r}\theta\right]\right\} + (1-\rho)\left\{\mathrm{K} + \sum_{j=1}^{n}\frac{\alpha_j}{r}\ln w_j + \frac{1}{r}\ln Y - \frac{\alpha_k}{r}\ln k + \sum_{i=1}^{l}\frac{\alpha_i}{r}\theta\right\} + \varepsilon_C,$$

with

(28)
$$\mathrm{H}_C = \xi\mathrm{K} - \frac{\sum_{i=1}^{l}\alpha_i/r}{\tau + \sum_{i=1}^{l}\alpha_i/r}\ln\frac{\sum_{i=1}^{l}\alpha_i/r}{\tau},$$

(29)
$$\xi = \frac{\tau}{\tau + \sum_{i=1}^{l}\alpha_i/r}.$$

The homogeneity property of degree one in input prices allows the parameters of the cost frontier to be identified. Once again, under a cost plus regime, the effort of the producer is nil and the structural cost frontier is similar to the one usually found in the literature. Now, if the effort of the producer is strictly positive, more structure should be imposed on the frontier in order to properly identify the endogenous part of the global inefficiency term and distinguish the effects of moral hazard from the effects of adverse selection.

We propose now an application of these propositions, with data on the French urban transport industry.

2. **Application**

The regulation of urban transport in France serves as an illustration of the usefulness of our preceding discussion. First, we specify the cost frontier. Then, the estimation procedure and the results are presented.

We use a database that has been created in the early 1980s. It assembles the results of an annual survey conducted by the Centre d'Etude et de Recherche du Transport Urbain (CERTU, Lyon) with the support of the Groupement des Autorités Responsables du Transport (GART, Paris), a nationwide trade organization that gathers most of the local authorities in charge of a urban transport network. For our study, we have selected all urban areas of more than 100,000 inhabitants for a purpose of homogeneity. However, the sample does not include the largest networks of France, i.e., Paris, Lyon and Marseilles, as they are not covered by the survey. The result is that the panel data set covers fifty-nine different urban transport networks over the period 1985-1993.

2.1. ***The cost frontier***

The empirical work involves fitting the stochastic cost function presented in Equation (27) to this panel data set of French urban transport networks. The level of supply to be offered is set by the authority and is then considered as exogenous by the operator. Hence, the use of a cost frontier to describe the technology is appropriate. We assume that the production process requires four inputs. These inputs are labor *L*, materials *M*, soft capital *S* and capital *k*. From Equation (27), the cost function to be estimated is

$$\begin{aligned}(30)\qquad \ln C = {} & \rho\left\{\beta_0{}' + \xi\left[\beta_L \ln w_L + \beta_M \ln w_M + \beta_S \ln w_S + \beta_Y \ln Y + \beta_k \ln k + \beta_L \theta\right]\right\} + \\ & (1-\rho)\left\{\ln \beta_0 + \beta_L \ln w_L + \beta_M \ln w_M + \beta_S \ln w_S + \beta_Y \ln Y + \beta_k \ln k + \beta_L \theta\right\} + \varepsilon_C .\end{aligned}$$

Note that only labor is considered as the potential candidate for primal inefficiency and asymmetries of information.[4] The parameters in (30) are all functions of the production frontier parameters. Thus, $\ln\beta_0 = \mathrm{K}$, $\beta_j = \alpha_j / r$, $j = L, M, S$, $\beta_y = 1/r$, $\beta_k = -\alpha_k / r$, $\xi = \tau / (\tau + \beta_L)$, and $\beta_0' = \ln\beta_0 + \beta_L(\ln\tau - \ln\beta_L - \ln\beta_0)/(\tau + \beta_L)$.

Estimating the Cobb-Douglas cost function requires measures on the level of operating costs, the quantity of output and capital and the input prices. Total costs C are defined as the sum of labor, materials and soft capital costs. Output Y is measured by the number of seat-kilometers, i.e., the number of seats available in all components of rolling stock times the total number of kilometers traveled on all routes. Capital k, which plays the role of a fixed input in our short-run cost function includes rolling stock. Since the authority owns capital, the operators do not incur capital costs. The average wage rate w_L is obtained by dividing total labor costs by the annual number of employees. Materials include fuel, spares and repairs. As the number of buses actually used mainly determines these expenditures, one derives an average price of materials w_M by dividing material expenditures by the number of vehicles. Soft capital includes commercial vehicles, computer service and office supplies. These charges are induced by the activity of network management. By dividing investment charges by the number of customer trips per year, one obtains the price w_S of managing single consumer travel. Summary statistics on the variables used in the analysis are given in Table I.

2.2. ***Estimation***

For a network *i* at period *t*, the stochastic cost function can be stated from Equation (30) as

(31) $$\ln C_{it} = C(Y_{it}, k_{it}, w_{it}, \theta_i, \rho_{it}, \beta) + \varepsilon_{c,it}.$$

We choose to estimate the parameters of the cost function by maximum likelihood. At this point, we need to make some assumptions about the distribution of the unobservable terms θ and ε

in (31). We assume that these terms are independent. The usual assumption found in the previous literature on parametric stochastic frontiers is to consider the second error component to be distributed as $N(0,\sigma_\varepsilon^2)$. We do the same here. We use a Beta function with scale parameters μ and ν to describe the distribution of θ. This choice is dictated by two considerations. First, in view of the relation between the efficient and actual levels of labor defined by Equation (5), the labor inefficiency parameter is conveniently defined as a percentage. This is readily obtained since the Beta density is defined over the interval $[0,1]$. As a matter of fact, choosing a beta density is an adequate normalization that does not impose strong restrictions. Second, the advantage of using it is that no specific shape is specified a priori. The distribution is symmetric if $\mu = \nu$, asymmetric otherwise and it can be hump-shaped or U-shaped. The likelihood of a data point conditional to θ is

(32) $$L_{it}(\theta) = L\left(\ln C_{it} \middle| Y_{it}, k_{it}, w_{it}, \theta_i, \rho_{it}, \beta, \sigma_\varepsilon, \mu, \nu\right) = \phi\left[\frac{\varepsilon_{c,it}}{\sigma_\varepsilon^2}\middle|\theta_i\right],$$

where $\phi(.)$ denotes the normal density function. Since the variable θ_i is unobservable, only the unconditional likelihood can be computed, i.e.,

(33) $$L_{it} = \int_0^1 L_{it}(u) u^{\nu-1}(1-u)^{\mu-1} \frac{\Gamma(\nu+\mu)}{\Gamma(\nu)\Gamma(\mu)} du .$$

where $\Gamma(.)$ is the gamma function. Assuming that the observations are independent, the log-likelihood function for our sample is just the sum of all individual log-likelihood functions obtained from Equation (33).

*2.3. **Technical inefficiency, effort and cost distortions***

Table II presents two estimation procedures. The first one provides an estimation of the structural cost frontier (30), which includes implicitly and explicitly the asymmetric information parameters e and θ. The second one proposes the standard fixed-effects model of a simple Cobb-Douglas cost frontier with no asymmetric information parameters

$$\ln C_{it} = \beta_i + \beta_L \ln w_{L,it} + \beta_M \ln w_{M,it} + \beta_S \ln w_{S,it} + \beta_Y \ln Y_{it} + \beta_k \ln k_{it} + \varepsilon_{c.it}, \tag{34}$$

with a firm specific effect β_i. This model is also referred as the least squares dummy variable model. In comparing the estimates of both specifications, note how the estimate of σ_ε falls in the asymmetric information model. Since (significant) coefficient estimates of both models are very similar, it is of interest to test which of the two specifications is the most appropriate. As the two models are non-nested, we use a test proposed by Vuong (1989). When the Vuong's statistic is less than 2 in absolute value, the test does not favor one model or the other. Otherwise, large positive or negative values favor one model or the other. The statistic of the asymmetric information model versus the fixed-effect one is 15. Thus, it strongly favors the structural approach presented in this paper.

The estimated distribution of θ has an exponential shape that is skewed to the right, which suggests that the operators are on average rather efficient (the probability to pick up a producer with a θ lower than 0.5 is greater than one half).

Estimates of individual efficiency parameter can also be recovered. As shown by Equations (30) and (31), the stochastic cost frontier has two random, unobservable and independent components θ_i and $\varepsilon_{c,it}$. In case of cost-plus regimes, the unobservable term is

$$u_{it} = \varepsilon_{c,it} + \beta_L \theta_i, \tag{35}$$

whereas, under fixed-price regimes, it is

(36) $$u_{it} = \varepsilon_{c,it} + \xi\beta_L\theta_i.$$

From a procedure initiated by Jondrow et al. (1982), one may recover an estimate for each θ_i from the values of residuals $\hat{u}_{it}$. To do so, we consider the conditional distribution of θ_i given u_{it}, i.e., by computing $\hat{\theta}_i = E(\theta_i / u_{it})$.

Once estimates for individual efficiency and effort levels have been recovered, we can evaluate the individual cost distortions above the stochastic frontier. Their expression is given by

(37) $$\exp[\beta_L(\theta - e)].$$

Table III lists the estimated technical inefficiency, effort levels and cost distortions over the frontier for the biggest networks included in our dataset. The other networks are available upon request. A distortion equal to 1.015, as in Toulouse for example, suggests that the observed operating costs are, on average, 1.5% above the frontier.

Consider Figure 1 where we present our set of fifty-nine networks ranked according to their cost distortions. Figure 1 provides for each network, the level of the inefficiency parameter and indicates the type of contract used to regulate it. Note that three groups of networks are easily detected. The first group with the lowest levels of cost distortion gathers sixteen networks, all of which are managed under a fixed-price contract. The next twenty ones can be collected in a second group as all of them (but four networks) are regulated through a cost-plus contract. Finally the last twenty networks are assembled in a third group, almost equally shared between the two types of contract. Concerning the third group, we just conclude that technical inefficiency is so high that even a highly incentive scheme, such as a fixed-price contract, cannot cure the problem. These results show that, because we account explicitly for the effect of each type of contract on the cost function and that our sample covers a large spectrum of existing networks, we are able to fully recover the distribution of the efficiency parameter.

Note that there are networks for which $e > \theta$. Since we did not impose $\theta > e$ in the course of the estimation, we obtained cases where the effort is slightly greater than inefficiency, implying negative cost distortions. In fact, these estimated negative cost distortions are not significantly different from 0. The estimated variances of the cost distortions are available upon request.

3. **Conclusion**

In this article, we have proposed an example of a regulated producer to work out structural frontiers and to shed light on an important problem that may arise when one estimates frontiers. The usual global inefficiency term is partly endogenous since it depends on the will of the producer. Failing to account for this potential endogeneity may lead to estimation errors.

We thus tried to provide some evidences on the reasons why it is useful to develop a structural analysis to well identify productive inefficiency. The global inefficiency has to distinguish the technical inefficiency that is exogenous and the endogenous effort that depends on the technological and regulatory conditions in a very specific way. Thus, the compensation for inefficiency in public or private firms may be searched in the system of incentives and institutional constraints.

Notes

1. In the Liebenstein's view, where the motivation is weak, firm management permits a considerable slack in the operations and does not seek cost improving methods. Since cost improving methods lead to internal costs, it appears that, in situations where there are no payments by result or competitive pressure, firm managers prefer to reduce the disutility of greater effort rather than decreasing operating costs. The literature on stochastic production frontiers has mentioned the existence of productive effort as a potential candidate for the appearance of inefficiencies (Aigner, Lovell and Schmidt, 1977, suggested: *Any deviation [from the frontier] is the result of factors under the firm control, such as technical and economic efficiency, the will and effort of the producer and his employees*). However it has not proposed any structural framework that takes into account the influence of incentives.

2. The introduction of asymmetric information in a production frontier model can be motivated by the simple fact that contracts for inputs are generally incomplete. In the case of labor for instance, it is exceedingly difficult to spell out all elements of performance in a labor contract. Neither individuals work as hard nor do they search for information as they could. Hence, technical inefficiency and effort might not be revealed by the producer to an observer who does not take part in the production process.

3. We could assume that all variable inputs are affected by inefficiencies, i.e., $l = n$. In this case, the global inefficiency term $(\theta - e)\sum_{i=1}^{l}\alpha_i \big/ r$ in Equations (10) and (13) simplifies to $(\theta - e)$. Assuming that $l < n$ allows us *(i)* to propose a general framework that can be easily simplified to the case where $l = n$, *(ii)* to shed light on the emergence of allocative efficiency, which does occur if all variable inputs are affected by the same technical inefficiency, and, *(iii)* to be coherent with the application proposed in Section 2. See also Footnote 4 for more details.

4. In the public transit industry, the network operator is better informed on labor inefficiency than the regulator. Note that labor costs represent more than 60% of total costs in this industry. Bus drivers play a decisive and acute role in operating the network, especially with respect to the flexibility and punctuality of operations in peak periods. First, bus drivers permanently meet the end users. Their

behavior vis-à-vis the customers may perceptibly affect the quality of service during high peak periods. Indeed, the driver has to perform several tasks at the same time, selling tickets, monitoring the passengers' up-and-down movements, managing the use of bus seats and space. Clearly, these tasks are much harder to accomplish in period of traffic congestion. Moreover, drivers have to deal with social and security problems, particularly in areas where the underprivileged population is large. There is an additional feature worth to be mentioned. The network structure may affect the driving conditions. On a same network, each bus route has its own characteristics of traffic lanes, route length, road access that complicates the evaluation of drivers' skills. All these remarks have the same implication: Appraising efficiency by just looking at the observed quantity of physical input is more difficult in the case of labor than in the cases of materials and soft capital whose consumption can be more easily observed and monitored by the regulator. As a result, we distinguish between observed and efficient labor forces, i.e., the physical amount of labor forces that is the source of cost distortions and is observable by the authority, and, the efficient level of labor required to produce the output. Since the behavior of drivers is the source of cost distortions, we assume that managers spend time and effort in monitoring drivers, providing them with training programs, solving potential conflicts, etc. Both labor technical inefficiency and cost reducing activity are unobservable to the regulator and to the econometrician.

References

Afriat, S.N., 1972, “Efficiency Estimation of Production Functions,” *International Economic Review,* 13, 568-598.

Aigner, D.J., C.A.K. Lovell and P. Schmidt, 1977, “Formulation and Estimation of Stochastic Frontier Production Models,” *Journal of Econometrics,* 6, 21-37.

Baron, D., and R. Myerson, 1982, “Regulating a Monopolist with Unknown Costs,” *Econometrica,* 50, 911-930.

Cornwell, C., P. Schmidt and R. Sickles, 1990, “Production Frontiers with Cross-Sectional and Time-Series Variations in Efficiency Levels,” *Journal of Econometrics,* 46, 185-200.

Farrel M.J., 1957, “The Measurement of Productive Efficiency,” *Journal of Royal Statistical Society A,* 120, 253-281.

Ivaldi, M., I. Perrigne and M. Simioni, 1994, “Productive Efficiency of French Grain Producers: A Latent Variable Model,” *Journal of productivity analysis,* 5, 287-299.

Jondrow, J., C.A. Lovell, I.S. Materov and P. Schmidt, 1982, “On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model,” *Journal of Econometrics,* 19, 233-238.

Laffont, J.J. and J. Tirole, 1986, “Using Cost Information to Regulate Firms,” *Journal of Political Economy,* 64, 614-641.

Leibenstein. H., 1966, “Allocative Efficiency Versus "X-Efficiency,” *American Economic Review,* 56, 392-415.

Loeb, M., and W. Magat., 1979, “A Decentralized Method of Utility Regulation,” *Journal of Law and Economics,* 22, 399-404.

Meeusen, W., and J. Van den Broeck, 1977, “Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error,” *International Economic Review,* 18, 435-444.

Schmidt, P., and R.C. Sickles, 1984, “Production Frontiers and Panel Data,” *Journal of Business & Economic Statistics,* 2, 367-374.

Vuong, Q., 1989, “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses,” *Econometrica*, 57, 307-334.

Table I: Descriptive statistics on the cost structure

Variable	Mean	Standard deviation
Total cost (10^3 FF)	117500.000	137731.000
Wage (10^3 FF)	174.940	28.384
Material price (10^3 FF)	26.311	31.386
Soft capital price (10^3 FF)	8.000	5.918
Capital (# vehicles)	143	134
Production (10^3seat-kilometers)	151302.680	367805.920
Labor share	0.573	0.128
Material share	0.296	0.117
Soft capital share	0.129	0.078

Table II: Estimation results

	Standard model		Asymmetric information model	
Parameters	Estimation	Standard error	Estimation	Standard error
β_O			0.3068	0.150
β_L	0.4285	0.041	0.4491	0.048
β_S	0.1027	0.011	0.0824	0.006
β_Y	0.0400	0.037	0.1825	0.022
β_K	0.7063	0.092	0.7010	0.048
$\ln \tau$			4.2827	0.257
ν			0.5931	0.035
μ			1.8007	0.287
σ_ε	0.1300	0.012	0.0834	0.007
Log-likelihood	0.549		0.594	
Sample size	531		531	

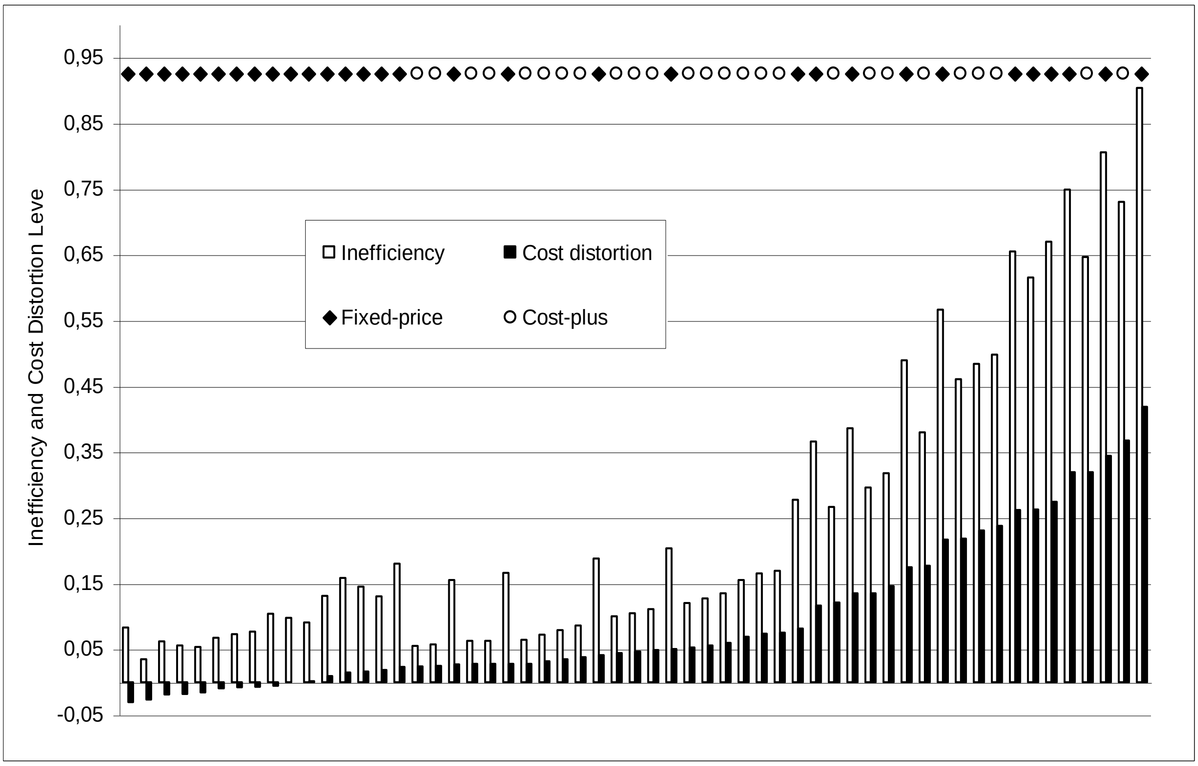
Note: The fifty-nine firm specific constant terms β_i of the standard model are not reported here. They are available upon request.

Table III: Technical inefficiency, effort level and cost distortion for some networks

Network	Technical inefficiency	Effort	Distortion
Aix	0.067	0.089	0.990
Besançon	0.318	0.000	1.153
Bordeaux	0.086	0.000	1.039
Caen	0.749	0.103	1.337
Cannes	0.646	0.000	1.337
Clermont	0.155	0.000	1.072
Dijon	0.120	0.000	1.055
Grenoble	0.083	0.114	0.986
Le Havre	0.266	0.000	1.127
Lille	0.180	0.126	1.024
Montpellier	0.131	0.110	1.009
Nantes	0.104	0.117	0.994
Nice	0.489	0.113	1.184
Nîmes	0.035	0.097	0.972
Rennes	0.484	0.000	1.243
Strasbourg	0.806	0.117	1.363
Toulon	0.064	0.000	1.029
Toulouse	0.158	0.124	1.015
Valence	0.111	0.000	1.051

Note: Under cost-plus regulation, the effort level is equal to zero.

Figure 1: Inefficiency and regulatory schemes



Note: To each network are associated three data: The inefficiency level (white bar), the cost distortion (black bar) and the type of contracts (a black diamond refers to a fixed-price contract and an empty circle indicates a cost-plus contract.